

Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft

Sax U¹, Viezens F¹, Mohammed Y¹, Hartung M², Weisbecker A³, Falkner J³, Bart J³, Knoch, T A⁴, Rienhoff O¹

¹ Abteilung Medizinische Informatik, Universität Göttingen

² Institut für Informatik, Universität Leipzig

³ Fraunhofer IAO, Stuttgart

⁴ Kirchhoff Institut für Physik, Universität Heidelberg und Erasmus Medical Center, Rotterdam

Summary

Summary (Sax)

Einleitung

Nach der Sequenzierung des menschlichen Genoms¹ hat sich das Feld der genomischen Medizin ausgeprägt. Durch zusätzliche Tests der genomischen Prädisposition von Patienten, lässt sich deren Behandlung in vielen Fällen individualisieren. Beispielsweise wurde bei der Behandlung von Rectumkarzinomen klassische immer die gleiche 5x5 Fu Radiochemotherapie durchgeführt. Durch die gezielte Untersuchung von Response und Toxizität und deren genomischen Markern kann man zukünftig non-Repondern bzw. Patienten mit einer guten Prognose die zweite Radiochemotherapie ersparen².

Für eine derartige Personalisierung zugunsten der Patienten ist jedoch eine gute Datenbasis eine Schlüsselkomponente. Während die genomischen Untersuchungen hochstandardisiert durchgeführt werden, ist dies auf der Phänotypseite (= tatsächliche Ausprägung beim Patienten) oft nicht der Fall. Diese Daten werden zunehmend nicht mehr in Papierform sondern in elektronischen Patientenakten abgelegt. Diese jedoch oft in durch Institutionsgrenzen vorgegebenen Datensilos.

Ein Weg, diese Grenzen aufzuweichen und alle relevanten medizinischen Daten zu einem Patienten zu sammeln ist die Persönliche Gesundheitsakte (PHR), die auch in den freiwilligen Anwendungen der eGK³ geführt wird, sowie in USA bis in wenigen Jahren für alle Bürger eingeführt werden soll⁴.

Um aufwändige Untersuchungen in der genomischen Medizin oder in der Bildverarbeitung durchführen zu können, sind Speicher- und Computingressourcen bzw. spezielle Algorithmen notwendig, die im Rahmen einer PHR oder in einem Krankenhaus meist nicht kurzfristig zur Verfügung stehen.

Diese Services können im Rahmen der Virtualisierung von Rechner, Speicher- und Servicestrukturen zur Verfügung gestellt werden. Derzeit untersucht das BMBF-geförderte

Projekt MediGRID⁵ im Rahmen der D-Grid-Initiative⁶ die Rahmenbedingungen für einen Grid-Einsatz in der Medizin.

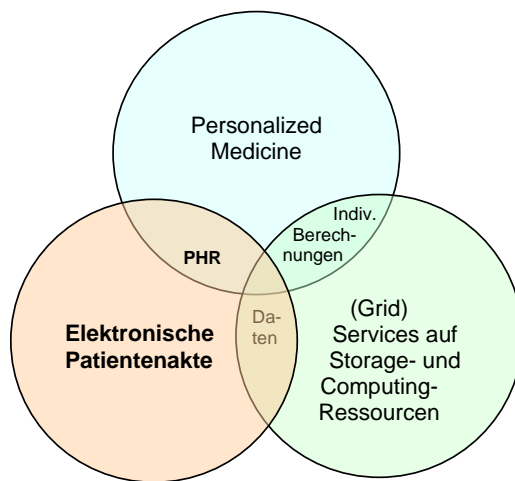


Abbildung 1: Venn-Diagramm des Überlapps von Personalisierter Medizin, Elektronischer Patientenakte und Grid-Services. Im Überlapp zwischen Personalized Medicine und Elektronischer Patientenakte ist die Persönliche Gesundheitsakte (Personal Health Record, PHR) zu finden.

Neue Möglichkeiten mit Grid-Ressourcen und Methoden im Grid

Zur Steigerung der Qualität der Patientenversorgung, bei gleichzeitiger Begrenzung der Kosten, ist es notwendig, die Prozesse sowohl bei den Leistungserbringern als auch bei den Kostenträgern zu optimieren und den Patienten in den Mittelpunkt zu stellen. Dazu werden IT-Lösungen benötigt, die die sektorübergreifende, patientenzentrierte Integration der Informationen aus den medizinischen und administrativen Prozessen unter Berücksichtigung der Datensicherheit und des Datenschutzes gewährleisten. Die Einführung der elektronischen Gesundheitskarte bildet die Basis für den Aufbau einer Telematikinfrastruktur. Zum zentralen Bestandteil wird sich dabei die elektronische Patientenakte entwickeln.

Mittelfristig werden die Grenzen zwischen ambulanter und stationärer Versorgung immer stärker verwischen. Aktuelle organisatorische Entwicklungen der integrierten Versorgung, wie beispielsweise Medizinische Versorgungszentren und Ärztenetze, können zukünftig daher nur durch eine leistungsfähige informationstechnische Unterstützung den Anforderungen der medizinischen Dokumentation, der Verwaltung der Patientendaten und der Qualitätssicherung genüge leisten. Gleichsam wird eine viel versprechende Ausgangslage für neue Dienstleistungen und Mehrwertapplikationen geschaffen.

Um den Entwicklungen in der medizinischen Versorgung und der Forschung gerecht zu werden ist eine leistungsfähige IT-Infrastruktur notwendig, die institutionsübergreifendes, vernetztes Arbeiten an verschiedenen Standorten sowie die gemeinsame Nutzung von Daten und Anwendungen unter Berücksichtigung des Datenschutzes und der Datensicherheit ermöglicht.

Grid Computing als neues Modell für die Nutzung von verteilten Ressourcen wie Rechenleistung, Daten und Software durch transparente Dienste erfüllt diese Anforderungen und eröffnet neue Möglichkeiten für die bedarfsgerechte Bereitstellung von Ressourcen.

Im Rahmen der deutschen e-Science-Initiative D-Grid wird in dem Verbundprojekt »MediGRID - Ressourcenfusion für Medizin und Lebenswissenschaften« (www.medigrid.de) eine Grid-Infrastruktur für die interdisziplinäre, standortunabhängige Zusammenarbeit von Medizinem und Bioinformatikern geschaffen. (Das Projekt MediGRID wird vom Bundesministerium für Bildung und Forschung unter dem Förderkennzeichen 01AK803A-H gefördert.) Derzeit sind Anwendungen aus den Bereichen Biomedizin, medizinische Bildverarbeitung und klinische Studien in MediGRID verfügbar. Neben den Anwendungsszenarien Bildverarbeitung, bioinformatische und klinische Forschung vervollständigenden Arbeiten in den Bereichen Middleware, Ressourcenfusion, Ontologie und eScience das Projekt bezüglich softwaretechnischer, semantischer, rechtlicher und internationaler Fragestellungen.

Unter Nutzung von rechen- und datenintensiven Anwendungen im Grid eröffnen sich neue Möglichkeiten der Datenverarbeitung und der Auswertung und Verknüpfung von Daten. Durch eine engere Vernetzung medizinischer Großgeräte sowie Rechner und Datenspeicher können, komplexe Anwendungen in der Medizin und Biomedizin von überall her genutzt und unter Sicherstellung des Datenschutzes auf medizinische Daten und Forschungsdaten zugegriffen werden. Der Zugang zu den leistungsfähigen Ressourcen an den verschiedenen Standorten in MediGRID ist weltweit jedem Nutzer mit einem Internetanschluss und der notwendigen Berechtigung möglich. Mithilfe moderner Sicherheitstechnologien werden im MediGRID dabei strenge Datenschutzrichtlinien umgesetzt, um die Sicherheit des Zugriffs zu gewährleisten.

Der weltweite Zugriff auf Anwendungen und Daten des MediGRIDs erfolgt durch das MediGRID Portal. Berechtigte Nutzer können dadurch Daten und Dienste des MediGRIDs nutzen, in dem sie sich durch ein elektronisches Zertifikat am Portal identifizieren und eine Zugangsberechtigung zu Anwendungen und Daten besitzen. Das Portal erlaubt es dem Benutzer sich auf seinen Aufgabenstellung in der medizinischen und biomedizinischen Forschung zu konzentrieren und sich nicht um die einzelnen, im Grid vorhandenen Ressourcen und deren Nutzung kümmern zu müssen. Um dies zu erreichen wurde mittels Fragebögen und Workshops bei den Nutzern und Ressourcenprovidern eine Anforderungsanalyse durchgeführt. Unter Einsatz des benutzerzentrierten Designs, bei dem der Prozess der Entwicklung einer Anwendung so zu gestalten ist, dass zu jedem Zeitpunkt die Bedürfnisse des späteren Benutzers im Mittelpunkt stehen und der Portal Analyse und Design Methode (PADEM ®) des Fraunhofer IAO, wurde das Portal entworfen und erprobt. Das Portal ist mit dem GridSphere Portal Framework realisiert, einem Open Source basierenden Webportal. Die Integration der medizinischen und biomedizinischen Fachanwendungen erfolgt durch Webservices. Die Anwendungen wurden in Gridsphere Portlets integriert, in denen die Anwendungen ablaufen. Für die Einbindung verteilter Datenbanken wurde der Grid Standard OGSA-DAI (Open Grid Service Architecture – Data Access and Integration) genutzt. Neben den anwendungsspezifischen Portlets besteht auch die Möglichkeit eigene Abläufe festzulegen und als Workflows zu definieren, die graphisch als Petrinetze dargestellt werden. Die Benutzer sind so in der Lage sich die Dienste auszuwählen, die ihre Aufgabenstellung am besten erfüllen. Die Kommunikation der Dienste während des Ablaufes findet ebenfalls über Webservice-Schnittstellen statt. Der Ablauf der Bearbeitung kann durch die Petrinetze visualisiert werden und erlaubt das Monitoring der Abläufe.

Im Portal werden verschiedene Nutzerrollen unterschieden. Entsprechend der Rolle und den damit verbundenen Berechtigungen werden dem Benutzer über das Portal, die für ihn zugänglichen Dienste und Daten zur Verfügung gestellt. Dabei werden die folgenden Nutzerrollen unterschieden. Es gibt einen Gastzugang zu MediGRID, der interessierten Personen anhand von ausgewählten Demonstrationen MediGRID vorstellt. Hierfür ist keine Zugangsberechtigung notwendig. Der MediGRID Benutzer erhält entsprechend seiner Berechtigung Zugang zu den Ressourcen, das umfasst Rechner, Geräte, Anwendungen und Daten. Der MediGRID Entwickler hat die Möglichkeit seine Anwendungen zu integrieren und mittels der Grid Resource Definition Language (GRDL) zu beschreiben, so dass sie von anderen genutzt werden können. Der MediGRID Ressourcenadministrator ist für den Betrieb seiner im Grid bereitgestellten Ressourcen, meist Rechner, Datenspeicher und Geräte verantwortlich. Die Eigenschaften der Ressourcen werden ebenfalls mit der Grid Resource Definition Language (GRDL) beschrieben. Der MediGRID Portaladministrator ist neben der Verwaltung und dem Betrieb des Portals für das Benutzer- und Rechteverwaltung zuständig.

In eigen entwickelten Middleware-Komponenten werden dem Benutzer in MediGRID Dienste bereitgestellt mit denen er seine Aufgabenstellungen erfüllen kann. Die Grundlage dafür bietet eine serviceorientierte Architektur, die in unterschiedlichen Schichten die Grid Services für MediGRID bereitstellt und eine einfache Integration der Anwendungen ermöglicht.

Ausschlaggebend für die erfolgreiche Nutzung ist die Integration von fachspezifischen Anwendungen in die Grid-Infrastruktur und eine leicht zu bedienende Benutzungsschnittstelle, die den einfachen Zugriff zu den Grid Ressourcen gewährleistet. In MediGRID ist so eine Grid-Infrastruktur entstanden, die die Zusammenarbeit vereinfacht und die Leistungsfähigkeit erhöht.

Kommentar [u1]: Sehr stark kürzen, hier zählt die Vision. Mehr noch das Argument: warum Grid in der PHR?

Anwendungen

^{7, 8} **6 7 8 9. 10 11 12 13 14. 15 16**

Analyse von medizinischen Bilddaten im Grid.

Im Bereich der Analyse von medizinischem Bildmaterial bietet das Grid Möglichkeiten der individualisierten Diagnose und Therapieplanung, die durch andere Technologie nicht oder nur schwer umsetzbar sind. Zum einen ermöglicht der gridweite Zugriff auf medizinische Bilddatenbanken eine strukturierte Bildsuche nach vergleichbaren Fällen als Diagnoseunterstützung, insbesondere bei selteneren Krankheitsbildern, die nicht aus der täglichen Routine bekannt sind.

Desweiteren benötigt die Analyse von medizinischen Bildern oft hochspezialisierte Algorithmen. Der serviceorientierte Ansatz moderner Grids, bei dem aus einem weiten Angebot von Analysetools und Bildverarbeitungsalgorithmen die jeweils benötigten ausgewählt und genutzt werden können, ermöglicht so die individualisierte Diagnoseunterstützung, ohne dass die entsprechende Software oder notwendige Hardware von dem behandelnden Arzt dauerhaft erworben werden muss.

Die stetig zunehmenden Bilddatenmengen, die durch Volumenaufnahmen und verbesserte räumliche und zeitliche Auflösung der Bilder erzeugt werden, stellen kontinuierlich steigende Anforderungen an Speicherplatz und Rechenleistung. Diese Anforderungen können durch ein Grid besonders gut erfüllt werden, da es jederzeit für den Nutzer unsichtbar erweitert und modernisiert werden kann.

Das Modul Bildverarbeitung des MediGrid-Projektes implementiert die wesentlichen Methoden medizinischer Bildverarbeitung sowie Datenstrukturen und -organisation (PACS, Bilddatenbanken) im MediGrid. Dadurch werden rechen- und speicherintensive Verfahren zur Verbesserung von Diagnose und Therapie ressourcenschonend ermöglicht und der verteilten Anwendung interdisziplinär zugänglich gemacht. Im Rahmen des MediGRID-Projektes werden drei prototypische Anwendungsszenarien der medizinischen Bildverarbeitung implementiert.

Die funktionelle Magnetresonanztomographie ermöglicht durch die statistische Analyse über große Sätze von Hirnbilddaten die Zuordnung von Hirnbereichen zu verschiedenen Stimulationen. Die Analyse stellt hohe Anforderungen an die Rechenkapazität und kann durch Parallelisierung auf Einzelbildebene durch den Einsatz im Grid stark beschleunigt werden. Das Anwendungsszenario wird unter der Leitung des Instituts für medizinische Informatik der Universität Magdeburg (Prof. Bernarding) implementiert.

Im Rahmen des virtuellen Gefäßchirurgie werden aufwendige hämodynamische Simulationen durchgeführt, wobei die Gefäßgeometrie aus computertomographischen Bilddaten extrahiert wird und durch den Nutzer verändert werden kann, um zum Beispiel die Implantation eines Stents zu testen. Die Algorithmen werden vom Institut für Strömungsforschung der Universität Erlangen (Dr. Beronov) eingebracht.

Das dritte Anwendungsbeispiel implementiert die komplexe Bildverarbeitungskette der computergestützten Diagnose von Prostatabiopsien. Die Gewebeprobe, die bei Verdacht auf Prostatakrebs entnommen werden, werden anhand von zwei- und dreidimensionalen transrektalen Ultraschallbildern im Prostatavolumen lokalisiert und visualisiert. Neben dem Institut für Medizinische Informatik der Charité bringen eine Reihe von weiteren Forschergruppen (DKFZ Heidelberg, RWTH Aachen, GSI Münschen, LMTU München) ihre Algorithmen ein, so dass alle Komponenten dem aktuellen Stand der Forschung entsprechen. Der Workflow umfasst auch den Zugriff auf PACS-Systeme und Image Retrieval Systeme. Durch die enge Zusammenarbeit mit den anderen Modulen des MediGridprojektes ist eine Standardisierung gewährleistet, die die Zusammenführung der genotypischen Daten und multidimensionale phänotypische Daten mit der medizinischen Bildverarbeitung technisch vereinfacht.

Die prototypischen Anwendungen sind somit richtungsweisend für eine individualisierte und leicht verfügbare computergestützte Diagnose und Therapieplanung.

Anwendungen aus der Biomedizin

Eine Pilotapplikation aus MediGRID (Augustus) wird zur Genvorhersage für ganze Genome und Genomabschnitte verwendet. Eine DNA enthält in ihrem Bauplan viele verschiedene Abschnitte für die Bildung der Proteine.

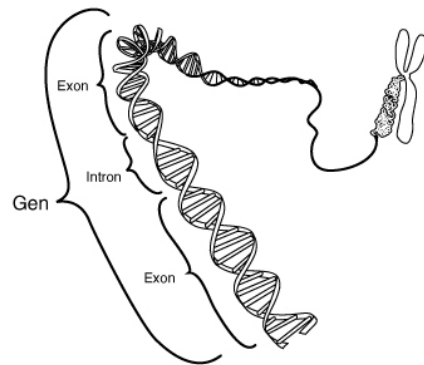


Abbildung 3.3.2: Genvorhersage für höhere Lebewesen

Dieses Programm kann mit hoher Genauigkeit Strukturen in den eukaryotischen genetischen Sequenzen vorhersagen. Die eingegebene DNA Sequenz wird ab initio oder mit generierten Hinweisen aus vorhandenen EST (expressed sequence tag)-Datenbanken gesucht. Bei der Nutzung der letzteren Variante werden bis zu 250 GB an Daten erzeugt und in der Ausgabe der Ergebnisse sind immerhin noch zwischen 5 und 50 MB an reinen Daten vorhanden. Durch eine stufenweise Analyse von überlappenden Abschnitten einer Sequenz ist eine Parallelisierung dieses Prozesses leicht umzusetzen. Die Anwender sind damit in der Lage verteilt rechnen zu können, mit einzelnen Instanzen eines Programms. Grid-Ressourcen erlauben zusätzlich das mehrfache Updates und die zentrale Ablage sehr großer EST-Datenbanken.

Im MediGRID-Portal wird diese Anwendung auf Grid-Ressourcen verteilt. Das Ergebnis ist nicht nur das gewohnte Text-File, es gibt auch eine hochauflösende Bildvariante für den Biomediziner, in der markante Abschnitte farblich gekennzeichnet sind.

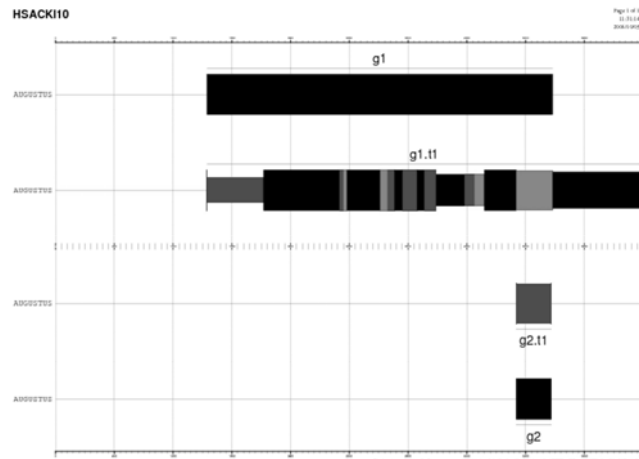


Abbildung 3.3.4: Grafische Darstellung des Ergebnisses aus dem AUGUSTUS-Portlet

Mit dieser Abbildung kann die Forschung die analysierte Sequenz mit den relevanten Abschnitten, die eine Genvorhersage aus einem oder mehreren Organismen beinhaltet, für weitere Vergleiche oder Experimente nutzen.

Kommentar [u2]: Neu, sax

Diskussion und Ausblick: EPA in einem Grid

Noch eine Stufe komplexer wird es bei medizinischen Grids, die in der klinischen Versorgung eingesetzt werden sollen. Wenn beispielsweise Krankenhäuser eine Grid-Infrastruktur für individuelle Therapieentscheidungen einsetzen wollen, müssen auf einmal nicht nur Rechenkapazitäten verwaltet werden, sondern es müssen auch medizinische Informationen aus verteilten Datenbanken zugänglich gemacht werden. Es müssen Algorithmen geschaffen werden, um diese Informationen sinnvoll auszuwerten. Es muss festgelegt werden, wer welche Daten nutzen darf und ob, wie und wem dafür Kosten in Rechnung gestellt werden dürfen. Und es gibt plötzlich ganz andere Anforderungen an den Datenschutz (siehe Kasten auf Seite XY). In all diesen Bereichen leistet das Projekt MediGRID Pionierarbeit. MediGRID ist das medizinische Teilprojekt der vom Bundesforschungsministerium seit 2005 mit insgesamt 15 Millionen Euro geförderten D-GRiD-Initiative. Es wird von den Göttinger Medizininformatikern und von der Telematikplattform für Medizinische Forschungsnetze (TMF) gemeinsam geleitet.

Literatur

1. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. Feb 16 2001;291(5507):1304-1351.
2. Ghadimi BM, Grade M, Difilippantonio MJ, et al. Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy. *J Clin Oncol*. Mar 20 2005;23(9):1826-1838.
3. Gematik. Technische Spezifikationen eGK. <http://www.gematik.de/download-spezifikationen.aspx>. Accessed 16.03.2006.
4. Thompson TG, Brailer JB. The Decade of Health Information Technology: Delivering Consumer-centric and Information-rich Health Care: Framework for Strategic Action. <http://www.hhs.gov/onchit/framework/>. Accessed August 16, 2004.
5. MediGRID. The MediGRID project.
. Available at: <https://www.d-grid.de/index.php?id=42&L=1>. Accessed 30.10.2006.
6. D-Grid. www.d-grid.de. Accessed march 7,2006.
7. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res*. Jul 1 2006;34(Web Server issue):W435-439.
8. Stanke M, Tzvetkova A, Morgenstern B. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biol*. 2006;7 Suppl 1:S11 11-18.
6. Ernst, T.; Falkner, J.; Peter, K.; Steinke, T.; Weisbecker, A. (2006): Analyse und Spezifikation der MediGRID-Anwendungen, MediGRID-Deliverable, 31.07.2006
7. Falkner, J.; Steinke, T., Weisbecker, A. (2006): Grid-Computing. In: Sax, U.; Mohammed, Y. Viezens, F.; Rienhoff, O. (Hrsg.): Grid-Computing in der biomedizinischen Forschung: Datenschutz und Datensicherheit. München: Urban und Vogel, S. 44-55
- 8 ISO 13407 (1999): Human-centred design processes for interactive systems
- 9 Spath, D.; Hinderer, H. (Hrsg.) (2005): Marktübersicht Portalsoftware 2005. Stuttgart: Fraunhofer IRB
- 10 Falkner, J.; Weisbecker, A. (2006): Integration of Applications in MediGRID. To appear in the conference proceedings of the Cracow Grid Workshop, 18.10.2006.
- 11 Hoheisel, A., Ernst, T., Der, U. (2006): A Framework for Loosely Coupled Applications on Grid Environments. In: Cunha, J. C.; Rana, O. F. (Eds.): Grid Computing: Software Environments and Tools. Heidelberg: Springer
- 12 Globus Toolkit, www.globus.org, 2006
13. J. Basney, M. Humphrey, and V. Welch. The MyProxy Online Credential Repository. Software: Practice and Experience, Volume 35, Issue 9, July 2005, pages 801-816
- 14 Glen Drinkwater, MyProxy Upload Tool, CCLRS, <http://tiber.dl.ac.uk:8080/myproxy/>
- 15 GridSphere Portal Framework, www.gridsphere.org
- 16 Grid Portlets, <http://www.gridsphere.org/gridsphere/docs/gridportlets/docbook/AdminGuide/AdminGuide.html>